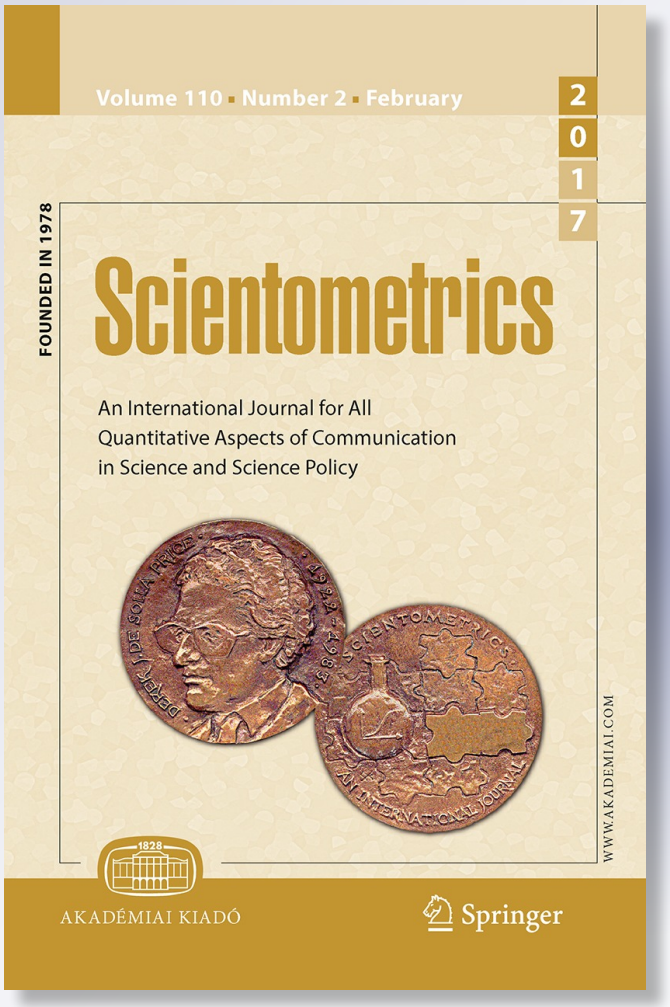
Highly cited publications in World War II: a bibliometric analysis

Yuh-Shan Ho & James Hartley

Scientometrics
An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science Policy

ISSN 0138-9130
Volume 110
Number 2

Scientometrics (2017) 110:1065-1075
DOI 10.1007/s11192-016-2199-4



Springer

Scientometrics (2017) 110:1065–1075
DOI 10.1007/s11192-016-2199-4

Highly cited publications in World War II: a bibliometric analysis

Yuh-Shan Ho[1] · **James Hartley**[2]

Received: 14 November 2016 / Published online: 3 December 2016

Abstract What are the characteristics of scientific papers published in World War II, and what papers from World War II, if any, are highly cited today? This paper reports that 3767 publications from World War II have been cited at least 100 times since 1939–1945. The data show that the publication rates of scientific papers declined during World War II only to increase rapidly after it. The USA was the most prolific source of scientific publications during the war, and Harvard University was the most dominant institute. In addition, there were five 'Sleeping Beauties', that is papers that were published but rarely cited during the war but came into prominence at a much later date.

Keywords Nobel laureates · Sleeping beauties · Web of Science · World War II · SCI-EXPANDED

Introduction

Figures 1 and 2 below show that publication rates decline during periods of war. However, little is known little about the effects of war on the remaining publications. In this paper we used the Science Citation Index to examine some of the properties of highly cited papers initially published during Word War II (1939–1945). Such highly cited articles are of particular importance because a high citation count is indicative of high impact in the research community, and highly cited articles provide interesting and useful insights into publishing over time (Smith 2008). Furthermore, high citation rates also reflect the impact of published works on the international research community.

✉ Yuh-Shan Ho
ysho@asia.edu.tw

[1] Trend Research Centre, Asia University, Taichung 41354, Taiwan

[2] School of Psychology, Keele University, Staffordshire ST5 5BG, UK

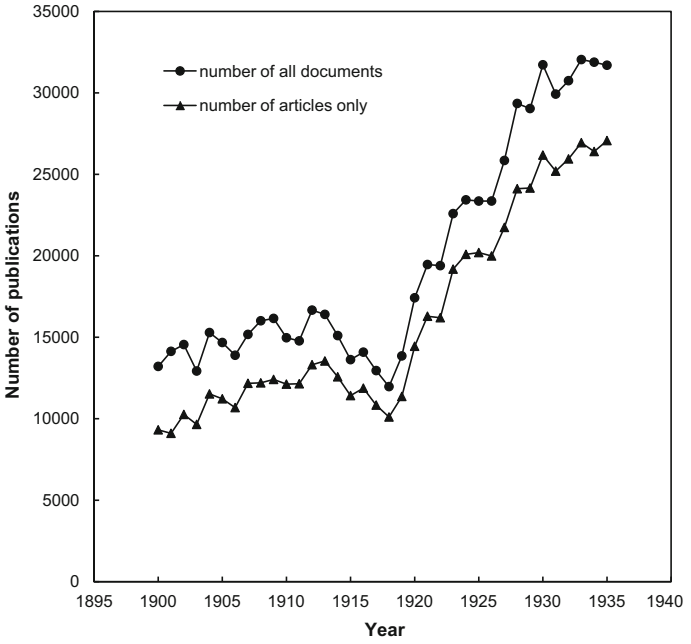


Fig. 1 Publications in World War I

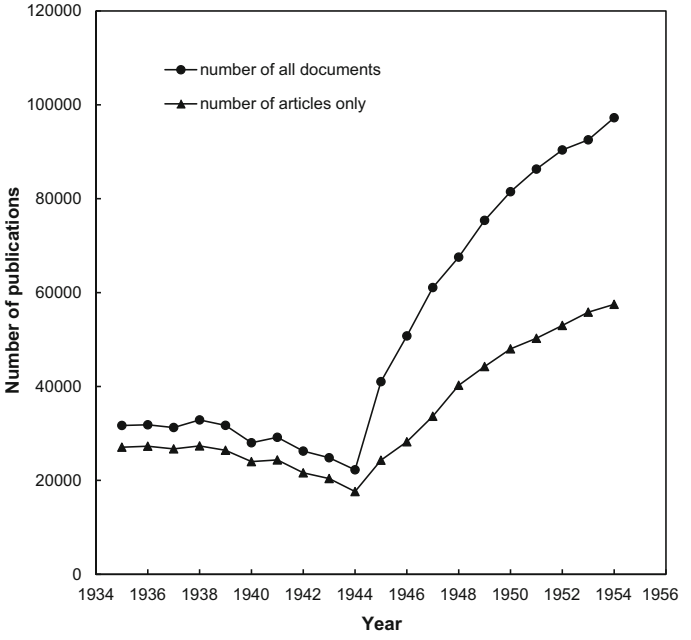


Fig. 2 Publications in World War II

In recent years, we and our colleagues have used the total number of citations since publication (TC_{year}) to evaluate highly cited articles (Chuang et al. 2011; Wang et al. 2011; Ho and Hartley 2016). The advantage of this measure is that it is invariant in comparison with the index of citations from the Web of Science Core Collection which has to be updated from time to time (Fu et al. 2012).

In this paper we report on some of the properties of the highly cited papers that were published during World War II.

Methodology

We used the Science Citation Index Expanded (SCI-EXPANDED) of Web of Science Core Collection (Thomson Reuters) to collect the bibliographic data. In 2015, this indexed 8864 journals across 177 subject categories in SCIE. In this particular study we searched for documents published between 1939 and 1945. Initially, we found 203,230 documents (as of June 13th, 2016). We then selected from these those papers that had been cited at least 100 times ($TC_{2015} \geq 100$). This yielded 3767 highly cited publications and we used the data derived from these publications for further analyses. In addition, we also used the number of citations in 2015 (C_{2015})—the most recent year then available—to characterize these highly cited papers (Ho 2013).

Results and discussion

It is clear from Figs. 1 and 2 presented above that similar publication trends were found during and after World War I and World War II respectively, with a decline during the war years and an increase after them (especially after World War II). In this study we analyzed those articles that were published in World War II that had had at least 100 citations in Web of Science Core Collection since publication to the end of 2015 ($TC_{2015} \geq 100$).

Document types

We found seven document types (Table 1). Ninety-four percent of these were articles followed distantly by reviews (4.5% of 3767 documents). Articles had the highest citations per publication (CPP) with 257 which is 3.2 times than reviews (CPP = 238). We chose articles in this bibliometric analysis because they contain description of complete researches and results (Ho et al. 2010).

Table 1 Document type distributions

Document type	TP (%)	TC_{2015}	CPP
Article	3523 (94)	904,749	257
Review	168 (4.5)	39,900	238
Letter	64 (1.7)	14,378	225
Editorial material	5 (0.13)	964	193
Note	4 (0.11)	693	173
Meeting abstract	2 (0.053)	840	420
Discussion	1 (0.027)	191	191

TP, total number of highly cited publications; TC_{2015}, the total number of citations from its date of publication to the end of 2015; CPP, citations per publication (TC_{2015}/TP)

Language of publication

96% of the articles were published in English, followed by German (3.5%), French (0.31%), and one each in Dutch, Rumanian, and Russian. The non-English article with the highest number of citations ($TC_{2015} = 1199$) was "Insulation and crystallization of the fermenting process of Enolase" (Warburg and Christian 1942).

Publication year

Figure 3 shows the number of articles published in World War II and the number of citations per publication by year. Of these, a total of 3523 articles (2.2% of 158,660 articles) were classified as highly cited articles.

Highly cited articles

3523 highly cited articles were published in 261 journals in 90 Web of Science categories in SCIE. Of these 261 journals: 63 (24%) contained only one highly cited article; 26 (10%) contained two; 22 (8.4%) contained three; and 22 (8.4%) contained four.

In total 2209 articles were published in 128 journals that had impact factors in 2015 and 1314 articles were published in 133 journals which did not.

Within the 90 Web of Science categories in SCIE, 31 categories (34%) generated 1–5 highly cited articles in World War II, 21 categories (23%) generated 31–100 articles, 17 categories (19%) generated 11–30 articles 11 categories (12%) generated 6–10 articles, and 10 categories (11%) generated more than 100 articles.

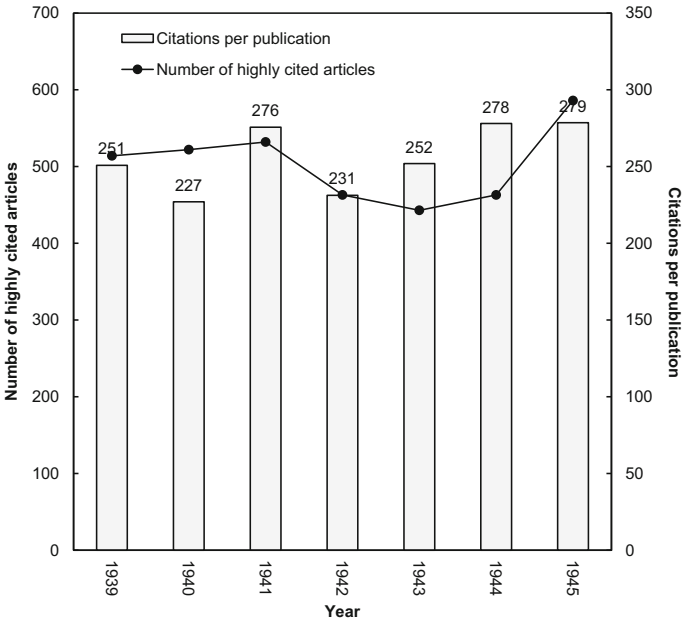


Fig. 3 Number of articles published in the World War II and citations per publication by years

Table 2 Top 12 Web of Science categories (TP ≥ 100)

Web of Science category	TP	%	Nos. journals
Multidisciplinary chemistry	391	11	163
Biochemistry and molecular biology	388	11	289
Physiology	275	7.8	83
General and internal medicine	255	7.2	151
Multidisciplinary physics	235	6.7	79
Research and experimental medicine	230	6.5	124
Multidisciplinary sciences	193	5.5	63
Neurosciences	179	5.1	256
Physical chemistry	143	4.1	144
Pathology	126	3.6	78
Immunology	100	2.8	150
Mathematics	100	2.8	312

TP, total number of highly cited articles in the World War II

Table 2 shows the 12 Web of Science categories each with at least 100 highly cited articles.

Publication performances: countries, institutions, and authors

To evaluate the publications of countries and institutions, 1920 articles (54% of the 3523 highly cited articles) with author information were analyzed, following the procedures advocated by Ho and Kahn (2014). Among the articles with this information, 1900 (99%) were country independent and only 20 (1.0%) involved international collaborations.

Table 3 lists the top ten countries that published highly cited articles in World War II. Five indicators were used as shown: total articles, independent articles, collaborative articles, first-author articles, and single-author articles. The USA convincingly took the first place for all of the indicators shown followed, distantly by the UK and the USSR. The USSR was also ranked 3rd for country independent and single author articles and,

Table 3 Characteristics of the top ten contributing countries in the World War II

Country	TP	TP *R* (%)	IP *R* (%)	CP *R* (%)	FP *R* (%)	SP *R* (%)
USA	1794	1 (93)	1 (93)	1 (90)	1 (93)	1 (90)
UK	30	2 (1.6)	2 (1.3)	2 (25)	2 (1.4)	2 (2.6)
Sweden	18	3 (0.94)	3 (0.79)	3 (15)	3 (0.89)	4 (1.4)
Canada	16	4 (0.83)	3 (0.79)	7 (5.0)	4 (0.83)	5 (1.2)
USSR	15	5 (0.78)	3 (0.79)	N/A	5 (0.78)	3 (1.6)
Holland	10	6 (0.52)	6 (0.42)	4 (10)	6 (0.52)	6 (0.54)
Denmark	6	7 (0.31)	7 (0.26)	7 (5.0)	7 (0.26)	6 (0.54)
Germany	6	7 (0.31)	7 (0.26)	7 (5.0)	7 (0.26)	9 (0.41)
China	4	9 (0.21)	13 (0.11)	4 (10)	11 (0.16)	9 (0.41)
Hungary	4	9 (0.21)	11 (0.16)	7 (5.0)	11 (0.16)	6 (0.54)

TP, total number of highly cited articles; IP, single country highly cited articles; CP, internationally collaborative highly cited articles; FP, first author highly cited articles; SP, single author highly cited articles; R, rank; %, percentage in each of TP, IP, CP, FP, and SP; N/A, not available

surprisingly, 12 of 20 internationally collaborative articles were published by single authors with affiliations in two countries.

Thirty-seven institutions published at least ten highly cited articles and the Academy of Sciences of the USSR was the only institution not located in USA that published more than ten articles. Table 4 lists the top ten institutions with more than 40 publications. Harvard University dominates the list (11%) followed by Columbia University (8.0%) and the University of Chicago (5.5%). In World War II Harvard University not only provided the most highly cited articles but also the most institutionally independent articles, inter-institutional collaborative articles, first author articles, and single author articles.

In total 4337 authors were involved in 3523 highly cited articles published during World War II. Of these: 3170 (73%) contributed one article, 675 (16%) contributed two, 243 (5.6%) contributed three, 103 (2.4%) contributed four, and 146 (3.4%) contributed five or more. Ninety-six of the authors (2.2% of 4337) won a Nobel Prize, including 40 in physiology or medicine, 31 in chemistry, and 25 in physics. In total Nobel laureates published 184 highly cited articles as first authors.

The most prolific author was P.J. Flory who published 13 highly cited articles in World War II, including the most first author, corresponding author, and single author publications. Flory won the Nobel Prize in chemistry in 1974 for his achievements in the physical chemistry of macromolecules—both theoretical and experimental.

The lifespan of the most cited articles in World War II

As shown in Table 5, only 21 articles were cited at least 100 times by 2015, and this number represents 0.60% of the highly cited articles published during World War II. Five of these 21 articles were published in *Journal of Chemical Physics*, and two in *Proceedings of the National Academy of Sciences of the United States of America* and *Physical Review* respectively. Sixteen articles were published by a single author, four by two authors, and one by four authors.

Table 4 Characteristics of the ten most productive institutions

Institution	TP	TP *R* (%)	IP *R* (%)	CP *R* (%)	FP *R* (%)	SP *R* (%)
Harvard University, USA	213	1 (11)	1 (9.2)	1 (23)	1 (9.6)	1 (11)
Columbia University, USA	153	2 (8.0)	2 (6.8)	2 (15)	2 (7.6)	2 (6.0)
University of Chicago, USA	106	3 (5.5)	3 (5.8)	6 (3.7)	3 (5.4)	2 (6.0)
University of Calif, USA	73	4 (3.8)	4 (3.8)	6 (3.7)	4 (3.5)	7 (2.2)
Yale University, USA	69	5 (3.6)	5 (3.6)	10 (3.3)	5 (3.3)	4 (4.4)
Cornell University, USA	59	6 (3.1)	6 (2.8)	5 (4.8)	6 (2.7)	5 (3.1)
Johns Hopkins University, USA	49	7 (2.6)	7 (2.4)	6 (3.7)	8 (2.3)	11 (1.6)
University of Pennsylvania, USA	48	8 (2.5)	9 (2.3)	6 (3.7)	7 (2.4)	6 (2.3)
University Wisconsin, USA	42	9 (2.2)	7 (2.4)	31 (1.1)	9 (2.1)	15 (1.4)
Caltech, USA	40	10 (2.1)	10 (1.9)	11 (3.0)	10 (1.9)	9 (1.9)

TP, total number of highly cited articles; IP, single institute highly cited articles; CP, inter-institutionally collaborative highly cited articles; FP, first author highly cited articles; SP, single author highly cited articles; R, rank; %, percentage in each of TP, IP, CP, FP, and SP

Table 5 Top 21 cited articles in 2015

Rank (C_{2015})	Rank (TC_{2015})	References
1 (584)	5 (4890)	Cassie and Baxter (1944)
2 (420)	40 (2018)	Patterson (1939)
3 (352)	8 (4303)	Wilcoxon (1945)
4 (337)	2 (6917)	Avrami (1939)
5 (330)	44 (1807)	Mann (1945)
6 (327)	16 (3193)	Dice (1945)
7 (294)	4 (5927)	Cole and Cole (1941)
8 (254)	24 (2537)	Maslow (1943)
9 (238)	13 (3413)	Murnaghan (1944)
10 (221)	15 (3197)	Biot (1941)
11 (212)	9 (4161)	Avrami (1941)
12 (191)	3 (5938)	Kramers (1940)
13 (188)	37 (2098)	Archie (1942)
14 (172)	22 (2803)	Wright (1943)
15 (168)	26 (2513)	Brunauer et al. (1940)
16 (142)	1 (10,845)	Nelson (1944)
17 (134)	73 (1294)	Hyers (1941)
17 (134)	34 (2203)	Bray and Kurtz (1945)
19 (131)	41 (2008)	Flory and Rehner (1943)
20 (112)	11 (3617)	Onsager (1944)
21 (100)	30 (2435)	Flory (1942)

C_{2015}, the total number of citations in 2015 from Web of Science Core Collection; TC_{2015}, the total number of citations from its date of publication to the end of 2015

The article by Cassie and Baxter (1944) had the highest number of citations by 2015 ($C_{2015} = 584$), and Avrami, of Columbia University, was the only author who published two articles with over 100 citations (Avrami 1939, 1941). An article by Nelson from the University of Cincinnati published in 1944 had the highest overall score of 10,845 citations ($TC_{2015} = 10{,}845$) and this paper still received 142 citations in 2015 ($C_{2015} = 142$). Two articles were published by Nobel laureates Flory and Rehner (1943) and Onsager (1944).

Figure 4 provides a picture of the citation lives of the top eight articles based on TC_{2015}. Four of these eight articles were also ranked among the top eight articles cited in 2015 (C_{2015}). These were the articles by Cassie and Baxter (1944) ranked 1st in C_{2015}, by Wilcoxon (1945) ranked 3rd, by Avrami (1939) ranked 4th, and by Cole and Cole (1941) ranked 7th. Other highly cited articles, for example Bratton and Marshall (1939) and Barker and Summerson (1941), have had a low impact in recent years, and it is widely accepted that the impact of a highly cited article might not always be high (Fu et al. 2012).

High impact sleeping beauties in World War II

Citation analyses can draw attention to papers that are suddenly discovered—or even "rediscovered"—after several years of dormancy (Garfield 1980). Gregor Mendel's (1866) "Versuche über pflanzenhybriden" is a typical example. This paper did not pass without notice, but its significance was not appreciated for over 30 years (Garfield 1980). van Raan (2004) discusses three main variables here:

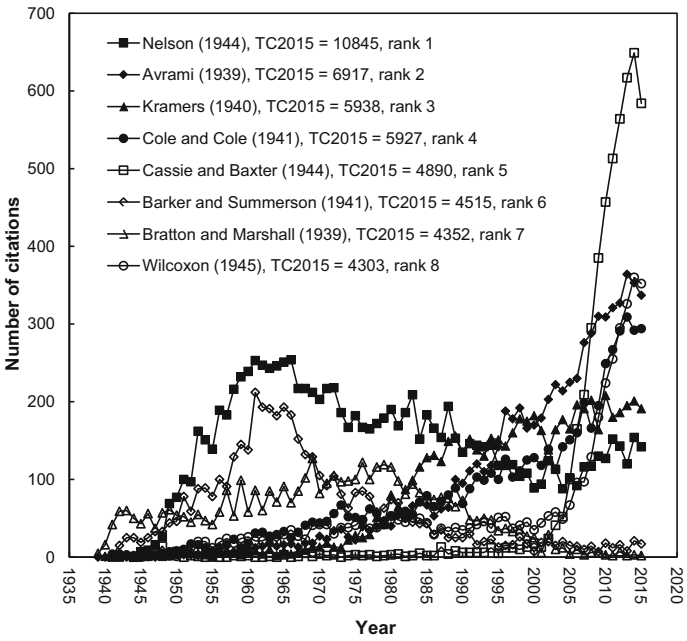


Fig. 4 Citation lives of the top eight articles with the $TC_{2015} > 4300$

Table 6 Six high impact sleeping beauties in the World War II ($C_{2015} \geq 100$)

C_{2015}	TC_{2015}	L_D	L_{LD}	L_H	Article title	References
420	2018	69	70	1	The Scherrer formula for X-ray particle size determination	Patterson (1939)
134	1294	57	62	8	On the stability of the linear functional equation	Hyers (1941)
327	3193	31	44	16	Measures of the amount of ecologic association between species	Dice (1945)
238	3413	26	33	28	The compressibility of media under extreme pressures	Murnaghan (1944)
221	3197	25	32	33	General theory of three-dimensional consolidation	Biot (1941)
212	4161	16	21	37	Granulation, phase change, and microstructure - Kinetics of phase change. III	Avrami (1941)

TC_{2015}, the total citations since publication to the end of the last year (2015); C_{2015}, the total citations in recent year (the last year 2015) only; L_D, length of the deep sleep (year); L_{LD}: length of the less deep sleep (year); L_H, years to reach 100 annual citations after the less deep sleep (year)

1. the depth of sleep, where an article receives at most one citation on average per year (deep sleep), or between one to two citations per year during a specific period (less deep sleep);
2. the length of sleep—the duration of the above period; and
3. the intensity of the wakeup period: the number of citations per year for four years following the sleeping period.

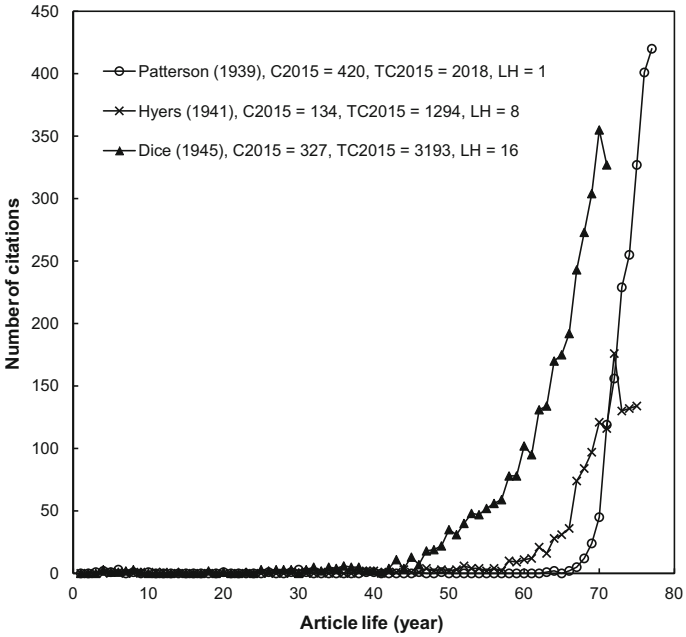


Fig. 5 The high impact sleeping beauty lifes of Patterson (1939), Hyers (1941), and Dice (1945)

Table 6 lists six high impact sleeping beauties from World War II and Fig. 5 shows typical citation curves for three of them. The life of the article by Patterson (1939) shown in Fig. 5 is typical of that of a high impact Sleeping Beauty. This particular article was in deep sleep for 40 years and then in less deep sleep for 20 more. It then spent only one year more to reach 100 annual citations. Similar curves are also shown in Fig. 5 for World War II articles by Hyers (1941) and Dice (1945).

Conclusions

In World War II, 3767 highly cited documents including 3523 articles were listed in SCI-EXPANDED, and English was the dominant language. The *Journal of the American Chemical Society*, the *Journal of Biological Chemistry*, and the *Physical Review* were the three most productive journals. Articles in the basic sciences such as chemistry, biochemistry, molecular biology, medicine, and physics were the main sources for these highly cited articles.

The USA contributed most highly cited articles and was also the most collaborative country, whilst the USSR was highly independent. The top 36 research sites in this period were in the USA, and Harvard University ranked top in six of the studied indicators. The USSR Academy of Sciences was ranked 37th in the total number of highly cited articles. Ninety-six of World War II authors won a Nobel Prize in physiology, medicine, chemistry, and physics and the most active authors were P.J. Flory and F. Albright. Six high impact sleeping beauties were published during World War II, and five of the top ten highly cited articles published during the war are still listed in the top ten articles of 2015. Thus, we

remain indebted to these scientists who worked during World War II, and it is a pleasure here to be able to recall their contributions.

References

Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the American Institute of Mining and Metallurgical Engineers, 146,* 54–61.

Avrami, M. (1939). Kinetics of phase change. I. General theory. *Journal of Chemical Physics, 7*(12), 1103–1112.

Avrami, M. (1941). Granulation, phase change, and microstructure: Kinetics of phase change. III. *Journal of Chemical Physics, 9*(2), 177–184.

Barker, S. B., & Summerson, W. H. (1941). The colorimetric determination of lactic acid in biological material. *Journal of Biological Chemistry, 138*(2), 535–554.

Biot, M. A. (1941). General theory of three-dimensional consolidation. *Journal of Applied Physics, 12*(2), 155–164.

Bratton, A. C., & Marshall, E. K. (1939). A new coupling component for sulfanilamide determination. *Journal of Biological Chemistry, 128*(2), 537–550.

Bray, R. H., & Kurtz, L. T. (1945). Determination of total, organic, and available forms of phosphorus in soils. *Soil Science, 59*(1), 39–45.

Brunauer, S., Deming, L. S., Deming, W. E., & Teller, R. (1940). On a theory of the van der Waals adsorption of gases. *Journal of the American Chemical Society, 62*(7), 1723–1732.

Cassie, A. B. D., & Baxter, S. (1944). Wettability of porous surfaces. *Transactions of the Faraday Society, 40,* 546–550.

Chuang, K. Y., Wang, M. H., & Ho, Y. S. (2011). High-impact papers presented in the subject category of water resources in the Essential Science Indicators database of the Institute for Scientific Information. *Scientometrics, 87*(3), 551–562.

Cole, K. S., & Cole, R. H. (1941). Dispersion and absorption in dielectrics I. Alternating current characteristics. *Journal of Chemical Physics, 9*(4), 341–351.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology, 26*(3), 297–302.

Flory, P. J. (1942). Thermodynamics of high polymer solutions. *Journal of Chemical Physics, 10*(1), 51–61.

Flory, P. J., & Rehner, J. (1943). Statistical mechanics of cross-linked polymer networks II swelling. *Journal of Chemical Physics, 11*(11), 521–526.

Fu, H. Z., Wang, M. H., & Ho, Y. S. (2012). The most frequently cited adsorption research articles in the Science Citation Index (Expanded). *Journal of Colloid and Interface Science, 379*(1), 148–156.

Garfield, E. (1980). Premature discovery or delayed recognition—Why. *Current Contents, 21,* 5–10.

Ho, Y. S. (2013). The top-cited research works in the Science Citation Index Expanded. *Scientometrics, 94*(3), 1297–1312.

Ho, Y. S., & Hartley, J. (2016). Classic articles published by American scientists (1900–2014): A bibliometric analysis. *Current Science, 111*(7), 1156–1165.

Ho, Y. S., & Kahn, M. (2014). A bibliometric study of highly cited reviews in the Science Citation Index Expanded™. *Journal of the Association for Information Science and Technology, 65*(2), 372–385.

Ho, Y. S., Satoh, H., & Lin, S. Y. (2010). Japanese lung cancer research trends and performance in Science Citation Index. *Internal Medicine, 49*(20), 2219–2228.

Hyers, D. H. (1941). On the stability of the linear functional equation. *Proceedings of the National Academy of Sciences of the United States of America, 27*(4), 222–224.

Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica, 7*(4), 284–304.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica, 13*(3), 245–259.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370–396.

Mendel, G. (1866). Versuche über Plflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn, 4,* 1–47.

Murnaghan, F. D. (1944). The compressibility of media under extreme pressures. *Proceedings of the National Academy of Sciences of the United States of America, 30*(9), 244–247.

Nelson, N. (1944). A photometric adaptation of the Somogyi method for the determination of glucose. *Journal of Biological Chemistry, 153*(2), 375–380.

Onsager, L. (1944). Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Physical Review, 65*(3–4), 117–149.
Patterson, A. L. (1939). The Scherrer formula for X-ray particle size determination. *Physical Review, 56*(10), 978–982.
Smith, D. R. (2008). Citation indexing and highly cited articles in the Australian Veterinary Journal. *Australian Veterinary Journal, 86*(9), 337–339.
van Raan, A. F. J. (2004). Sleeping beauties in science. *Scientometrics, 59*(3), 467–472.
Wang, M. H., Fu, H. Z., & Ho, Y. S. (2011). Comparison of universities' scientific performance using bibliometric indicators. *Malaysian Journal of Library & Information Science, 16*(2), 1–19.
Warburg, O., & Christian, W. (1942). Insulation and crystalisation of the fermenting process of Enolase. *Biochemische Zeitschrift, 310*(6), 384–421.
Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83.
Wright, S. (1943). Isolation by distance. *Genetics, 28*(2), 114–138.